บทที่ 2
ทฤษฎีพื้นฐานที่ใช้ในโครงงาน

ในหัวข้อนี้จะกล่าวถึงทฤษฎีพื้นฐานต่างๆ ที่เกี่ยวข้องในการวิจัย และพื้นฐานของระบบ License Plate Recognition ซึ่งเนื้อหาในบทนี้จะกล่าวถึงเทคนิคการแปลงภาพสีเป็นภาพขาวดำ การทำ Segmentation ระบบรู้จำตัวอักษร Neural Network ซึ่งเนื้อหาทั้งหมดนี้จำเป็นสำหรับการศึกษาและประเมินประสิทธิภาพของระบบ License Plate Recognition

2.1 **Digital Images**

2.1.1) Binary Images

เป็นการจับภาพจากสิ่งแวดล้อม หรือ ทำสำเนาภาพจากเอกสารให้อยู่ในรูปแบบของอิเล็กทรอนิกส์ เช่น รูปถ่าย เอกสารที่เขียนด้วยมือ เอกสารพิมพ์ และพิมพ์เขียว เป็นต้น โดย Digital images จะอยู่ในรูปของแผ่นตารางโดยแต่ละช่องจะเป็นส่วนหนึ่งของภาพหรืออักษร เรียกแต่ละจุดหรือช่องนั้นว่า “pixel” แต่ละ pixel จะถูกกำหนดให้มีระดับของความเข้ม (สีดำ สีขาว สีเทาหรือสีอื่นๆ) ซึ่งแสดงให้อยู่ในรูปของ รหัส Binary (0 และ 1) แต่ละ pixel ก็จะแทนด้วย Binary digital (“bits”) จะถูกเก็บเป็นลำดับใน computer และโดยทั่วไปจะถูกลดขนาดลงด้วยวิธีการทางคณิตศาสตร์ (บีบอัดให้เล็กลง) แต่ละ bit จะถูกแปลและอ่านโดย computer ให้เป็นแบบ Analog ซึ่งเป็นรูปภาพ หรือ แผ่นพิมพ์

1	1	1	1	1	1	1	1	1	1
1	0	0	0	1	1	0	0	0	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	0	0	0	0	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	0	0	0	1	1	0	0	0	1
1	1	1	1	1	1	1	1	1	1

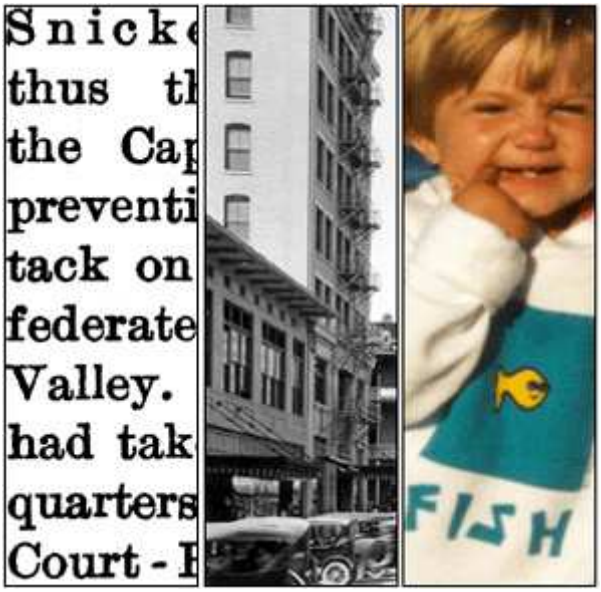
รูปที่2.1 ภาพแสดงแต่ละ Pixel ถูกกำหนดให้มีค่าเป็น 0 กับ 1 โดย 0 แสดงเป็นสีดำและ 1 แสดงเป็นสีขาว

2.1.2) Grayscale image

เป็นการเรียงของ pixel ที่ใช้ข้อมูลแบบ multiple bits อยู่ในช่วงระหว่าง 2- 8 bit หรือมากกว่านั้น ตัวอย่าง 2-bit image จะมีได้ 4 รูปแบบสี คือ 00 01 10 และ 11 ถ้า 00 คือสีดำ และ 11 คือสีขาว แต่ 01 คือดำเทา และ 10 คือเทาสว่าง bit depth คือ 2 แต่จำนวน tone จะเป็นค่า 2^2 หรือ 4 ที่ 8 bit = (2^8) = 256 tone ที่ต่างกันเป็นตัวกำหนดค่า pixel

2.1.3) Color image

แบบทั่วไปนั้นจะมีค่า bit depth อยู่ในช่วง 8 – 24 bit หรือมากกว่า ภาพที่มี 24 bit นั้นคือ bit จะถูกแบ่งออกเป็น 3 กลุ่ม 8 สำหรับสีแดง 8 สำหรับสีเขียว 8 สำหรับสีนำเงิน สีทั้งหมดจะถูกรวมกันเพื่อแสดงสีอื่นๆ 24-bit image สามารถแสดงค่าสีได้ถึง 16.7 ล้านสี (2^{24}) สำหรับ scanner นั้นได้เพิ่ม จำนวน bitในการจับภาพเอกสารเป็น 10 bit หรือมากกว่านั้น เพราะบ่อยครั้งหลังจากจับภาพเอกสารที่ 8 bit จะมี noise รวมเข้าไปด้วย และเพื่อให้ภาพที่ออกนั้นมีความเหมือนใกล้เคียงกับที่มนุษย์ต้องการ



รูปที่2.2 จากซ้ายไปขวาภาพ 1-bit 2สี, ภาพ 8-bit grayscale และ ภาพสีแบบ 24-bit

การคำนวณเลข Binary เพื่อให้ทราบจำนวนของสีที่แสดงในรูป Bit depth

1 bit (2^1) = 2 สี

2 bits (2^2) = 4 สี

3 bits (2^3) = 8 สี

4 bits (2^4) = 16 สี

8 bits (2^8) = 256 สี

16 bits (2^{16}) = 65,536 สี

24 bits (2^{24}) = 16.7 ล้าน สี

2.1.4) File size

คำนวณได้จาก การคูณกันของพื้นผิวของเอกสาร (สูง x กว้าง) ที่ถูกเลือก คูณกับ bit depth และdpi2 เนื่องจาก ขนาดของ file ภาพนั้นจะต้องถูกแสดงในรูปของ byte ที่ซึ่งก็คือ 8 bits ดังนั้นจึงต้องหารด้วย 8

สูตร 1 สำหรับคำนวณหาขนาด File
ขนาด file = (height x width x bit depth x dpi^2) / 8

(2.1)

ตัวอย่างเช่น ภาพที่ได้มาจากกล้อง Digital มีขนาด 24 bit โดยมีจำนวน pixel dimension 2,048 x 3,072 ดังนั้นขนาด file จะเท่ากับ (2,048 x 3,072 x 24)/ 8 หรือมีค่าเท่ากับ 18,874,368 byte

สูตร 2 สำหรับคำนวณหาขนาด File
ขนาด file = (pixel dimensions x bit depth) / 8

(2.2)

File size naming convention: ด้วยเหตุที่ว่า file ภาพ digital ส่วนมากมีขนาดใหญ่มาก จำนวน byte จึงเป็นหน่วยที่ถูกนำมาใช้เพื่อแสดงจำนวนค่า ที่เพิ่มตั้งแต่ 2^{10} ขึ้นไป

1 Kilobyte (KB) = 1,024 bytes

1 Megabyte (MB) = 1,024 KB

1 Gigabyte (GB) = 1,024 MB

1 Terabyte (TB) = 1,024 GB

2.2 การแปลงภาพสีเป็นขาว - ดำด้วยวิธี Threshold

การแสดงผลของภาพที่มีความเข้มหลายระดับ เช่น ภาพ 16 ล้านสีบนอุปกรณ์ ที่มีการแสดงผลได้แค่ 2 ระดับ (ขาว-ดำ) นั้นจะต้องแปลงข้อมูลภาพให้เป็นภาพแบบ Binary (Binary Image) ซึ่งการสร้างภาพแบบ Binary มีกระบวนการแปลงภาพที่มีความเข้มหลายระดับ(Multilevel Image) ให้เป็นภาพที่มีความเข้มเพียง 2 ระดับ หรือ 1 บิต (bit) คือ 0 กับ 1โดย 0 แทน ด้วยจุดที่มีภาพสีขาว และ 1 แทนด้วยจุดที่มีภาพสีดำ

เมื่อผ่านขั้นตอนการแปลงภาพนี้แล้ว นำภาพที่ได้ไปแสดงผลยังอุปกรณ์เหล่านั้น ซึ่งจะเห็นได้ว่าการแปลงข้อมูลภาพหลายระดับเป็นภาพ Binary จะมีความจำเป็นและมีประโยชน์อย่างมากในการแสดงผล การแปลงเป็นภาพ Binary ยังมีประโยชน์อีกคือในการแปลงข้อมูลภาพเป็นภาพ Binary คือการลดเนื้อที่การเก็บข้อมูลภาพ จะใช้เนื้อที่การเก็บ 8 บิต หรือ 256 ระดับ เมื่อสร้างเป็นภาพ Binary แล้ว สามารถลดลงจากเดิมได้ถึง 8 เท่า นั่นคือ 1 จุดจะใช้เนื้อที่ในการเก็บทั้งหมดคือ 1 บิต ซึ่งก็คือทำภาพที่มีความละเอียดสูงหรือโทนสีมาก ให้เหลือความละเอียดน้อยลง การแสดงโทนสีก็จะน้อยลงตามคือเหลือขาวกับดำ

การทำภาพ Binary ทำได้โดยใช้เทคนิค Threshold (Thresholding Technique) คือการพิจารณาจุด pixel ในภาพว่าจุดใดควรจะเป็นจุดขาว ที่มีค่าเท่ากับ 0 หรือจุดใดควรจะเป็นจุดที่มีค่าเท่ากับ 1 โดยจะทำการเปรียบเทียบค่าของแต่ละ pixel (f(x,y)) กับค่าคงที่ ที่เรียกว่าค่า Threshold (Threshold Value) เทคนิคนี้มีใช้กันมากในกรณีมีความแตกต่างระหว่างวัตถุ (Object) และพื้นหลัง (Background) ซึ่งเป็นลักษณะเดียวกับสีขาวของป้ายทะเบียนและสีดำของตัวอักษรบนป้ายทะเบียนค่าของ pixel ในภาพที่มีค่าน้อยกว่าค่าThresholdจะถูกกำหนดเป็น 1 (จุดดำ) และถ้าค่าของ pixel ใดๆในภาพที่มีค่ามากกว่าหรือเท่ากับค่า Threshold จะถูกกำหนดให้เป็น 0 (จุดขาว) ในการทำภาพ Binary โดยการทำ Threshold ให้ได้ภาพดีและคมชัดต้องเกิดจากการเลือกค่า Threshold ที่ถูกต้องและเหมาะสม

ถ้าเลือกค่า Threshold ไม่เหมาะสม เช่น ค่า Threshold ที่มากเกินไปหรือน้อยเกินไป ภาพที่ได้จะขาดความคมชัดหรือออาจะทำให้รายละเอียดของภาพขาดหายไปหรือภาพที่ได้อาจจะมืดเกินไปหรือสว่างเกินไปหรืออาจจะเป็นภาพที่มีสิ่งรบกวน (Noise) เกิดขึ้นทำให้ภาพผลลัพธ์ที่ได้ไม่ชัดเจนเท่าที่ควร ดังนั้นปัญหาในการสร้างภาพแบบ Binary คือ ทำอย่างไรจึงจะคำนวณค่า Threshold ที่เหมาะสมกับภาพ แต่ละภาพที่จะมาสร้างเป็นภาพแบบ Binary ซึ่งค่า Threshold สามารถคำนวณได้หลายวิธี แต่ละวิธีจะเหมาะกับการทำงานที่แตกต่างกัน

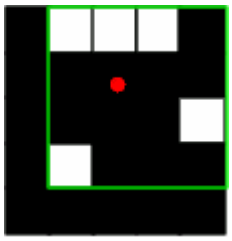
วิธีแรกเป็นการใช้ค่า Threshold จากการคำนวณ (Mid-Range Threshold Value) วิธีนี้จะเป็นการคำนวณโดยอัตโนมัติโดยไม่ต้องให้ผู้ใช้เป็นผู้กำหนดการหาค่า Threshold วิธีนี้ใช้หลักการคำนวณพื้นฐานทางสถิติในเรื่องของการหาค่ากลางหรือค่าเฉลี่ย (Mean) มาประยุกต์ใช้ค่า Threshold ที่คำนวณได้จากค่ากึ่งกลางที่อยู่ระหว่างค่าที่มี

ความเข้มสูงสุด (Maximum Level) และระดับความเข้มต่ำสุด (Minimum Level) ของข้อมูลอินพุท เมื่อทำการคำนวณค่า Threshold ได้แล้วก็จะสามารถนำค่า Threshold เป็นตัวกำหนดในการสร้างภาพ Binary ได้

วิธีที่สองเป็นการหาค่าThresholdโดยการกำหนดก่อนนำไปใช้ (Pre-assigned Threshold Value) วิธีนี้เป็นวิธีที่ง่ายในการคำนวณค่าโดยการกำหนดเองของผู้ใช้ซึ่งการกำหนดนี้จะขึ้นอยู่กับประสบการณ์ของผู้ใช้ โดยการเลือกค่าคงที่ที่เรียกว่าค่า Threshold นั้นค่าที่เลือกมานี้จะเป็นค่าที่อยู่ระหว่างค่าสูงสุดและค่าต่ำสุดของข้อมูลอินพุทแต่ละ pixel ของภาพ เช่น ภาพข้อมูลอินพุทมี Gray Level 256 ระดับจะมีค่าทั้งหมด 0-255 เมื่อเลือกค่า Threshold แล้วก็สามารถนำค่า Threshold เป็นตัวกำหนดในการสร้างภาพ Binary ได้

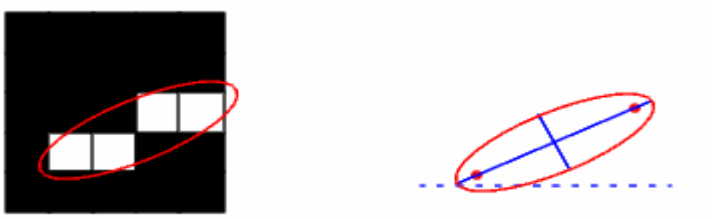
2.3 การแบ่งป้ายทะเบียนเป็นส่วนๆโดยวิธี Segmentation

เมื่อรู้ตำแหน่งของป้ายทะเบียนแล้ว เราจะเข้าสู่วิธีการแบ่งตัวอักษรเป็นส่วนๆ โดยวิธี Segmentation การทำวิธีนี้จะทำให้ตัวอักษรในป้ายทะเบียนถูกแบ่งออกเป็นส่วนๆ โดยจะทำการกำหนดขอบ Bounding Box ล้อมรอบส่วนของภาพที่มีขนาดใกล้กันจากนั้นระบุจุด Centroid ของรูปซึ่งจุดนี้คือจุดศูนย์มวลของรูปใน Bounding Box ไม่ใช่จุดศูนย์กลางของภาพ



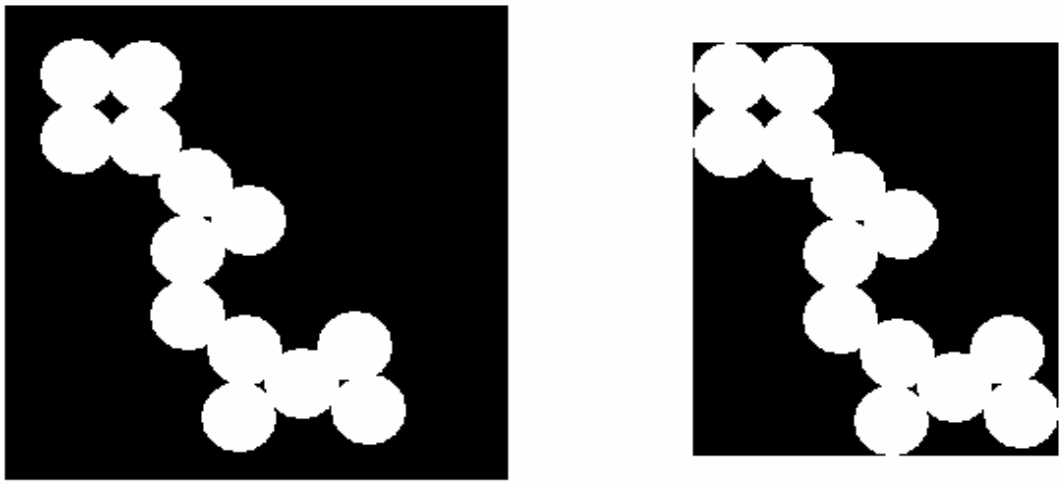
รูปที่2.3 แสดงขอบเขตของ Bounding Box ขอบสีเขียว กับจุด Centroid สีแดง

จากรูปจะเห็นจุด Centroid สีแดงอยู่ต่ำลงมาในแนว Vertical 2 ช่องในแกน Y และอยู่มาทาง Horizontal ด้านขวา1ช่องในแกน x ซึ่งจุดนั้นเองคือจุดศูนย์มวลของรูป Pixel สีขาว 5 รูปใน Bounding Box



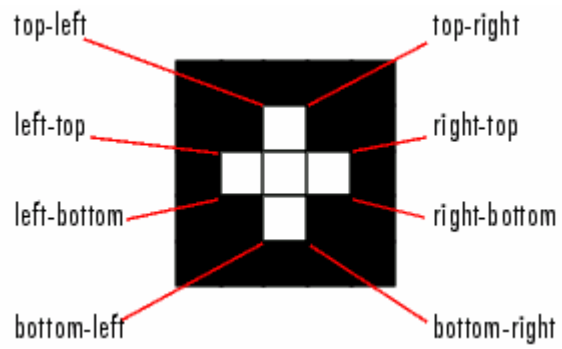
รูปที่2.4 แสดงขอบเขตของ Bounding Boxในแนวทะแยง

จากรูปจะเห็นวิธีการทำ Bounding Box ในแนวทแยงซึ่งจะทำการ Generate แกนแนวขวางสีฟ้าและกำหนดจุดระหว่างมุมสีแดงเพื่อใช้ในการทำเส้นทแยงมุมของแกนแนวขวาง



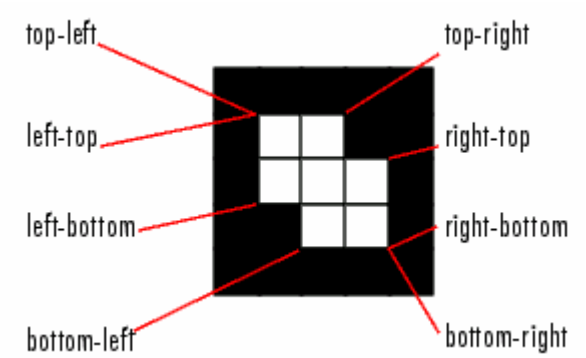
รูปที่2.5 แสดงภาพก่อนและหลังจากการทำ Segmentation

รูปทางซ้ายมือเป็นรูปก่อนการนำเข้ากระบวนการ Segmentation ส่วนรูปด้านขวาเป็นผลจากการทำงานผ่านกระบวนการ Segmentationซึ่งจะเห็นว่าขอบที่ไม่จำเป็นของตัวอักษรถูกตัดทิ้งไป



รูปที่2.6 แสดงวิธีแบ่ง Segment ของภาพในแนว Horizontal และแนว Vertical

จากรูปจะทำการแบ่งขอบของรูปออกเป็นส่วนใหญ่ๆ คือแนว Top , Bottom , Left , Right ซึ่งจะเป็นการแบ่งในแนวตั้งกับแนวนอนของรูป



รูปที่2.7 แสดงวิธีแบ่ง Segment ของภาพในแนวทแยง

จากรูปจะทำการแบ่งขอบของรูปออกเป็นส่วนใหญ่ๆ คือแนว Left , Right , Top , Bottom แต่จะเน้นที่มุมของรูปในแนวทแยงเพื่อต้องการกำหนดแกนของรูปในแนวทแยง